

# DEEP LEARNING METHOD FOR MEDICAL IMAGE ANALYSIS

*Elcin Huseyn,*

*Research Laboratory of Intelligent Control and Decision Making Systems in Industry and Economics,  
Azerbaijan State Oil and Industry University, Baku, Azerbaijan,  
ORCID ID: <https://orcid.org/0000-0001-5965-7419>*

**DOI:** [https://doi.org/10.31435/rsglobal\\_conf/30012021/7387](https://doi.org/10.31435/rsglobal_conf/30012021/7387)

**Abstract.** *With the continuous development and progress of medical imaging technology and computer technology, medical image analysis has become an indispensable tool and technical means in medical research, clinical disease diagnosis and treatment. In recent years, deep learning (DL), especially deep convolutional neural networks (CNN's), has rapidly developed into a research hotspot in medical image analysis. Diagnostic characteristics of the disease. This article first briefly describes the characteristics of medical image analysis. Second, it discusses the basic principles of deep learning and introduces the main deep learning technologies mainly used in medical image analysis. Then it discusses deep learning in medical image classification, detection, segmentation, registration, and retrieval. Domestic and foreign research status in various fields such as image generation and enhancement; finally, discuss the future challenges and strategies for inductive deep learning methods for medical image analysis.*

**Keywords:** *medical image analysis; deep learning; convolutional neural network*

**1. Introduction.** In the field of medical imaging, when a doctor or researcher performs quantitative analysis, real-time monitoring, and treatment planning on a specific internal tissue or organ, in order to make a correct treatment decision, it is usually necessary to understand some detailed information of this tissue and organ. Therefore, biomedical imaging has become an indispensable part in the diagnosis and treatment of diseases, and it is increasingly important.

Medical imaging technologies such as magnetic resonance image (MRI), positron emission tomography (PET), computer tomography (CT), cone beam CT, and 3D ultrasound imaging have been widely used for clinical examination, diagnosis, treatment and decision-making. How to make full use of artificial intelligence deep learning methods to analyze and process these large-scale medical image big data, provide scientific methods for the screening, diagnosis, and efficacy evaluation of various major diseases in clinical medicine, is a major science urgently to be solved in the field of medical image analysis Questions and key technologies for cutting-edge medical imaging.

Medical image analysis mainly used edge detection, texture features, morphological filtering, and construction of shape models and template matching. This type of analysis method is usually designed for a specific task and is called a manual custom design method. Deep learning is a data-driven analysis of tasks, which can automatically learn relevant model features and data characteristics from large-scale data sets for specific problems. Unlike explicit manual design models for specific problems, deep learning methods can implicitly automatically learn medical image features directly from data samples. The learning process is essentially an optimization problem solving process. Through learning, the model selects the correct features from the training data so that it can make the correct decision when testing new data. Therefore, deep learning plays a vital role in medical image analysis.

In recent years, deep learning has made significant progress, mainly due to the continuous improvement of computing power and the continuously increasing amount of available data, as well as the continuous improvement of deep learning models and algorithms. The essence is to build a multi-hidden layer machine learning model, use massive sample data to train, learn more accurate features, and finally improve the accuracy of classification or prediction [1]. Deep learning is a feature that learns hierarchical features from data, making it very suitable for discovering complex structures in high-dimensional data. At present, deep learning has been applied to image recognition, speech recognition, natural language processing, weather prediction, gene expression, and content recommendation. And other fields and various challenges.

The great success of deep learning in the field of computer vision has inspired many scholars at home and abroad to apply it to medical image analysis. In recent years, many experts have summarized, commented and discussed the research status and problems of deep learning in medical image analysis.

Recently, a review published on Medical Image Analysis summarized the research on deep learning in medical image classification, detection and segmentation, registration and retrieval, etc. [2].

## 2. Characteristics of Medical Image Analysis.

Medical image analysis has been widely used in clinically-assisted screening, diagnosis, grading, treatment decision-making and guidance, and efficacy evaluation of major diseases such as benign and malignant tumors, brain function and mental disorders, and cardiovascular and cerebrovascular diseases. Medical image classification and recognition, localization and detection, tissue and organ and lesion segmentation are the main application areas of current deep research methods for medical image analysis. There is a big difference between medical image analysis with different imaging principles and natural image analysis in the field of computer vision. So far, domestic and foreign scholars have carried out a series of deep learning research work mainly on medical image analysis tasks of different imaging principles such as MRI, CT, X-ray, ultrasound, PET, and pathological optical microscope. Therefore, this section mainly summarizes the analysis of these kinds of medical images.

### 2.1. Common Medical Images

**1) MRI image:** nuclear magnetic resonance image (MRI), which is a measure of the size of the magnetic resonance signal generated by the hydrogen atomic nucleus in human tissues and lesions under the action of an external strong magnetic field, and is received by a computer to an external nuclear magnetic resonance signal detector the information data is subjected to 3D image reconstruction. It can provide very clear human soft tissue anatomy and lesion images.

**2) CT image:** Computer tomography (CT) uses a precisely collimated X-ray beam to scan a section of a certain thickness of a part of the human body, and the X-ray transmitted through the section is received by a detector that rotates with the radiation beam, and finally, the computer reconstructs the 3D image of the corresponding section of the human body based on the X-ray signal data received by the detector. It has sub-millimeter-level spatial resolution, can provide clear human bone tissue anatomy and lesion images, and has been widely used in a variety of clinical disease examinations and auxiliary diagnosis.

**3) X-ray image:** Medical X-ray image is an electronic density measurement image of different tissues and lesions of the human body. X-ray based imaging includes 2D computer radiography, digital X-ray photography, digital subtraction angiography and mammography, and 3D spiral computer tomography, etc., and has been widely used in orthopedics, lungs, etc., Breast and cardiovascular diseases and other clinical diseases detection and auxiliary diagnosis, but 2DX ray images can not provide three-dimensional information about human tissues and organs and lesions.

**4) Ultrasound imaging:** Scan the human body with an ultrasound beam, and receive and process reflected signals to obtain images of internal organs. In recent years, ultrasound imaging technology has continued to develop, and new ultrasound imaging technologies such as 3D color ultrasound, ultrasound holography, intrabody ultrasound imaging, color Doppler imaging, and ultrasound biomicroscopy have appeared.

**5) PET image:** Positron emission tomography (PET) uses positron information emitted when a tracer labeled with a radioactive element such as F18 is decaying. Therefore, PET image is a measure of the radioactivity of the corresponding tracer and can provide tumor biology Information on characteristics (such as glucose metabolism, hypoxia, proliferation, etc.), and the standard intake value can be used to clinically determine the benign / malignant tumor. PET can provide more intuitive and accurate visual biological and radiobiological characteristics than CT and MRI.

**6) Pathological image:** refers to cutting out a certain size of diseased tissue, using hematoxylin and eosin (H & E) staining methods to make the sliced tissue into a pathological slide, and then using microscopic imaging technology to image microscopic cells and glands. By analyzing the pathological images, the cause, pathogenesis, and pathogenesis of the lesion can be explored to make a pathological diagnosis.

At present, the research objects of deep learning in clinical medical image analysis can be mainly divided into the above 6 types of medical images.

### 2.2. Main medical image analysis tasks

**1) Classification and recognition of medical images.** Clinicians need to use medical images to assist in diagnosing whether a human body has a lesion, and to quantify the severity of the lesion. Therefore, automatically identifying the lesion area and normal tissues and organs in the image is the basic task of medical image analysis.

**2) Medical image localization and detection.** The positioning of human tissues and organs and the location of lesions are very important pre-processing steps in clinical treatment planning and

intervention processes. The accuracy of positioning directly affects the effect of treatment. Image target positioning tasks not only need to identify specific targets in the image, but also determine their specific physical locations. Image target detection tasks need to identify all targets in the image and determine their physical locations and categories.

**3) Medical image segmentation task.** Image segmentation is to identify the internal voxels and external contours of the target area of interest in the image. It is a key task for clinical surgical image navigation and image-guided tumor radiotherapy.

### 3. Overview of deep learning methods

Machine learning algorithms are usually divided into supervised learning algorithms and unsupervised learning algorithms. In the supervised learning algorithm, not only the training data is input to the computer, but also the labels of the data are also passed as input to the computer. After the computer learns, it then passes new unknown data to it. Find out the probability that the data belongs to a certain result, and finally give you a result closest to the correct one. Because the computer has not only training data but also results (labels) during the learning process, the training results are often good. Different from supervised learning, the input data of unsupervised learning is only the original data and no labels, so the computer cannot accurately know which data belongs to which classes, and can only analyze the characteristics of the data by virtue of its powerful computing power to obtain certain results. Usually get some collections, the data in the collections are the same or similar in some characteristics.

#### 3.1 Neural Network

Neural network is an important machine learning technology in the field of machine learning and cognition. A mathematical or computational model that mimics the structure and function of a biological neural network and is used to estimate or approximate functions. The neural network is calculated by a large number of artificial neurons. In most cases, the artificial neural network can change the internal structure based on external information. It is an adaptive system, and in general, it has a learning function. Modern neural networks are a non-linear statistical data modeling tool. A typical neural network has the following three parts:

**1) Architecture:** Architecture specifies the variables in the network and their topological relationships. For example, the variables in a neural network can be the weights of neuron connections and the activities of the neurons.

**2) Stimulation function (Activity Rule):** Most neural network models have a short time scale dynamics rule to define how a neuron changes its own stimulation value based on the activities of other neurons. Excitation function is generally dependent on the weight in the network (i. e. the network parameter).

**3) Learning Rule:** The learning rule specifies how the weights in the network adjust over time. This is generally seen as one kind of long time scale dynamics of rules. In general, the value of the learning rule is dependent on the excitation of neurons. It may also depend on the weight of the target and current weight provided by the supervisor.

A simple neural network typically comprises three parts: an input layer, hidden layer and output layer. General node input layer and output layer is fixed, free to specify the hidden layer, Fig.1. is a schematic diagram of a simple neural network architecture, Figure circles represent neurons, arrows represent the flow of data, each of the above connections are corresponds to a different weight, the weight is the network by learning what you get.

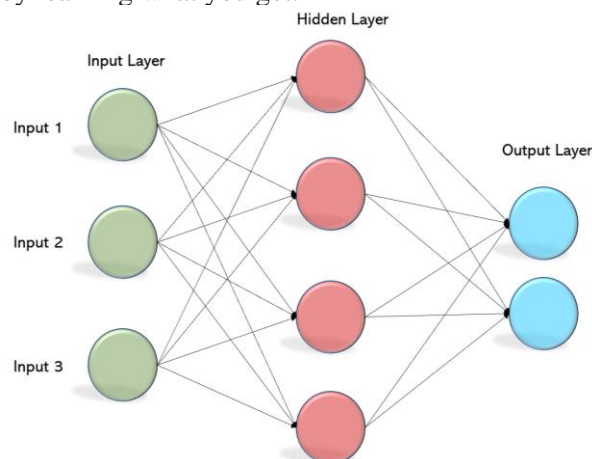


Fig.1. Neural Network Structure.

Each neuron contains the input, output and calculation model part 3, can be regarded as a neuron computing and storage unit, calculates neuronal function calculated input thereto. Temporary storage neurons is the calculation result, and passed to the next layer. Fig.2. is a typical neuron model, comprising three inputs, an output and a calculation function.

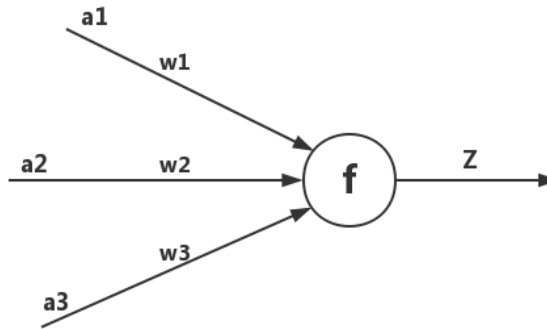


Fig.2. Neuron Structure Diagram.

Wherein a represents the input, W is the weight, indicates a connection to an arrow can be understood: in the beginning end, the signal transmission is still a size, intermediate end has a weighting parameter w, through this signal will be weighted into a w, so at the end of the connection, the magnitude of the signal becomes a w. The weighted signal in neuronal function f is obtained by summing the output Z, the following formula:

$$z = f(a_1 * w_1 + a_2 * w_2 + a_3 * w_3) \quad (1)$$

In fact, in order to achieve better results neurons, usually after an offset b for each input weighting plus will, at this time Z is calculated as follows:

$$z = f\left(\sum_{i=1}^n a_i * w_i + b\right) \quad (2)$$

Where the function f is also known as activation function, there is a common activation function sigmoid function, tanh function and ReLU functions and so on.

### 3.2. Neural Network Framework

**1) The depth of the neural network (DNN).** DNN deep layers are fully connected network, determines the number of layers of the neural network capability characterization data: using fewer neurons each fitting more complex functions. In 2006 Hinton with pre-training methods to alleviate the problem of local optima, will promote the hidden layer to layer 7, opened a wave of deep learning. Note that there is no fixed definition of "depth" here. In the field of speech recognition, 4 layers are considered to be relatively deep. In the image recognition field, 20-layer models are common. 2016 proposed highway network and the depth of the residual learning (deep residual learning) to further eliminate the gradient disappears reached a hundred layers.

**2) Convolutional neural network (CNN).** Convolutional neural network changes a neural network structure, the structure is not fully connected, and greatly reduce the amount of network parameters; while further reducing the parameter sharing network parameters. It takes into account the spatial structure and local features, ideal for image processing, neural network convolution is currently the most widely used application in the field of medical image analysis. For example, when segmenting a medical image, CNN can simply classify each pixel in the image by extracting patches around specific pixels.

**3) Cycle neural network (RNN).** Given the wide applicability of deep learning in real-world tasks, it has attracted the attention of many technical experts, investors, and non-professionals. Although the most famous achievement of deep learning is to use feed-forward convolutional neural networks to solve computer vision problems, a small amount of public attention has been devoted to using recurrent neural networks to model temporal relationships. In ordinary fully connected networks or CNN's, the signals of neurons in each layer can only propagate upwards, and the processing of samples is independent at each moment, so it is also called feed-forward neural network. In the RNN,



the output of the neuron can directly affect itself at the next timestamp; that is, the input of the neuron at the  $i$ -th layer at time  $m$  includes the output of the  $i-1$  layer at that time and its own at Output. Based on this, a long-term short-term memory LSTM network was developed.

**4) Deep Belief Network (DBN).** Probability confidence network depth (deep belief networks, DBN) comprising multiple layers of hidden units is a generative model, a composite model can be considered as a simple multi-layer combination of learning models. As part of the depth of a pre-trained neural network, and provide initial weight, then using back-propagation algorithm or the other is determined as a means to tune the network.

Although we can classify various methods of deep learning, in the broadest sense, NN or DNN includes other various methods. In actual use, it is often a fusion of multiple structures.

#### **4. Application of deep learning in medical images**

##### *4.1 Medical Image Classification*

Medical image classification can be divided into image screening and target or lesion classification. Image screening is one of the earliest applications of deep learning in the field of medical image analysis. It refers to taking one or more examination images as input, predicting it through a trained model, and outputting a signal indicating whether it has a certain disease or severity. Hierarchical diagnostic variables. Image screening belongs to image-level classification. The deep learning models used to solve this task initially focused on SAE, DBN, and DBM networks and unsupervised pre-training methods. Research has focused on the analysis of neuroimaging, such as diagnosing Alzheimer's disease (AD) or mild cognitive impairment (MCI) through neuroimaging [90–92]. These algorithms usually use multi-modal images as input to extract complementary feature information in modalities such as MRI and PET. At present, CNN is gradually becoming the standard technology in image screening and classification, and its application is very extensive. For example, Arevalo et al. Proposed a feature learning framework for breast cancer diagnosis, using CNN to automatically learn distinguishing features and classify breast X-ray lesions [3]. Kooi et al compared manual design and automatic CNN feature extraction methods in traditional CAD, both of which were trained on a large dataset of about 45,000 breast X-ray images, and the results show that CNN is superior to traditional CAD system methods at low sensitivity, and in Both are equivalent at high sensitivity [4]. Spampinato et al. Used deep CNN to automatically evaluate bone age [5]. In addition, there are some work combining CNN and RNN. For example, Gao et al. Used CNN to extract low-level local feature information in the slit lamp image, combined with RNN to further extract high-level features, and classified nuclear cataracts [6].

The classification of targets or lesions can help doctors diagnose the disease, such as the classification of benign and malignant breast lesions. The processing process usually first identifies or marks specific areas through pre-processing methods, and then classifies specific areas or targets. Accurate classification requires not only local information on the appearance of the lesion, but also global context information of its location. CNN is also widely used in the classification of targets or lesions. Kawahara et al. Used multi-processing stream CNN to classify skin lesions, where each process processes images with different resolutions [7]. Jiao et al. Used CNN to extract deep features at different levels and improved the classification accuracy of breast cancer [8]. Tajbakhsh et al. Compared the end-to-end artificial neural networks of Massive-training artificial neural networks (MTANNs) and CNN for the task of detecting pulmonary nodules in CT images and distinguishing between benign and malignant pulmonary nodules. The performance of the network, and its experimental results show that the performance of MTANN is significantly higher than that of CNN with less training data [9].

##### *4.2 Medical image positioning and detection*

Accurately positioning specific biomarkers or anatomical structures in medical images is of great significance in clinical treatment, and is directly related to the quality of the treatment. Medical image localization often requires analysis of 3D voxel information. In order to use classic deep learning algorithms for 3D data processing, some methods treat 3D space as a combination of 2D orthogonal surfaces, so that the positioning task can be converted into a classification task and processed using a general deep learning framework. For example, Yang et al combined the information of three orthogonal directions CNN to identify the marker of the femoral end, and the marker 3D position is defined as the intersection of three 2D tiles [10]. Vos et al. Analyzed the volume of 3D CT into 2D form and identified the 3D rectangular bounding box of the target, and then located the anatomical area such as heart, aortic arc and descending aorta of interest [11].

The key to detecting an object of interest or lesion in a medical image is to classify each pixel. At present, most deep learning-based target detection systems use CNN to perform pixel classification tasks,

and then use some form of post-processing to obtain the targets. The Chinese University of Hong Kong Chen et al. Used multiple 2D deep features to approximate the features of 3D medical images, combined with SVM classifiers, to achieve the automatic detection of Cerebral microbleeds (CMBs) with Susceptibility weighted imaging (SWI) [12]. There is still a small amount of work using other deep learning methods to achieve the target of interest or lesion detection. For example, Shin et al. Applied SAE deep learning methods to detect abdominal organs on MRI images, first learning spatial features in an unsupervised manner, and then performing multiple based on points of interest. Organ detection [13].

#### 4.3 Medical Image Segmentation

The task of medical image segmentation is usually defined as identifying the contours or the set of voxels that make up the object of interest. It is the most common topic in papers applied in the field of medical image analysis. Segmentation of organs and their substructures in medical images can be used to quantitatively analyze clinical parameters related to volume and shape, such as the ventricular volume and contractile ejection rate of the heart. On the other hand, when using intelligent intensity-modulated radiotherapy to treat tumors, endangering organ delineation is one of the very important steps in making a radiotherapy plan. Deep learning is widely used in this task, mainly for: histopathological images and microscope image segmentation; brain tissue structure segmentation and heart ventricle segmentation.

Computer segmentation of image features from surgical and biopsy specimens can help predict the extent of disease aggressiveness for disease diagnosis and grading. The key component of these predictors is image features mined from histopathological images. The current methods of segmentation of most logarithmic histopathological images and microscope images are based on CNN. Many scholars have achieved very good segmentation results using block training networks. Ciresan et al. Took the lead in applying deep CNN to medical image segmentation, and they used a sliding window method to segment biological nerve membranes from electron microscope images [14]. Kumar et al. Used block-based CNN to perform nuclear segmentation on H & E-stained pathological images [15].

The volume morphology of brain tissue structure is closely related to many neurological diseases of the brain. Such as depression, Alzheimer's disease, schizophrenia and bipolar disorder. Therefore, the study of the anatomical structure of brain tissue structure by computer technology plays a very important role in medical research, clinical diagnosis and treatment. For example, Zhang et al. Used T1, T2, and FA (Fractional anisotropy) modal images as input, and used a deep CNN network to solve the challenging task of infant GM, WM, and CSF segmentation, so as to evaluate the pros and cons of infant brain development [16].

Segmenting the left ventricle from cardiac MRI data is one of the important steps in calculating clinical indicators such as ventricular volume and systolic ejection rate. Carneiro et al. Used DBN-based learning features to model the appearance of the left ventricle, and used a supervised learning model to automatically segment the left ventricle in the cardiac ultrasound image [17], while Avendi et al. Used SAE to learn the deep features to initially infer the shape of the left ventricle, combined with the deformation the model improves the accuracy and robustness of left ventricular segmentation [18].

From the perspective of deep learning application frameworks, most current image segmentation methods are based on CNN. Many scholars have achieved good segmentation results using image block-based training networks.

**Conclusions.** In summary, deep learning has the ability to automatically learn deeper, more discriminative features from data. It has been applied to multiple research areas of medical image analysis and has made breakthroughs. We noticed that in most literatures, the use of deep learning-related methods has demonstrated its leading level of performance, which has been proven by the results of several computing challenges in medical image analysis; secondly, cloud computing and multi-GPU high performance parallel computing Development has made it possible for deep learning to learn deep features from large amounts of medical image big data; Finally, the emergence of publicly accessible related medical image databases and multiple medical image segmentation challenge datasets have enabled deep learning-based segmentation algorithms to Validated.

We believe that with the continuous improvement of deep learning algorithms, with the development of high-performance parallel computing technology and the increasingly improved medical image quality and growing medical image labeling sample set, deep learning-based medical image analysis will make a big difference.

REFERENCES

1. Zhang Lei, Zhang Yi. Big data analysis by infinite deep neural networks. *Journal of Computer Research and Development*, 2016, 53 (1): 68-79.
2. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JAWM, van Ginneken B, Sánchez C. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 2017, 42 (9): 60-88.
3. Arevalo J, González FA, Ramos-Pollán R, Oliveira JL, Lopez MA G. Representation learning for mammography mass lesion classification with convolutional neural networks. *Computer Methods and Programs in Biomedicine*, 2016, 127: 248-257.
4. Kooi T, Litjens G, van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, den Heeten A, Karssemeijer N. Large scale deep learning for computer-aided detection of mammographic lesions. *Medical Image Analysis*, 2017, 35: 303-312.
5. Spampinato C, Palazzo S, Giordano D, Aldinucci M, Leonardi R. Deep learning for automated skeletal bone age assessment in X-ray images. *Medical Image Analysis*, 2016, 36: 41-51.
6. Gao XT, Lin S, Wong T Y. Automatic feature learning to grade nuclear cataracts based on deep learning. *IEEE Transactions on Biomedical Engineering*, 2015, 62 (11): 2693-2701.
7. Kawahara J, Hamarneh G. Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained layers. *International Workshop on Machine Learning in Medical Imaging*. Athens, Greece: Springer, 2016. 164-171.
8. Jiao ZC, Gao XB, Wang Y, Li J. A deep feature based framework for breast masses classification. *Neurocomputing*, 2016, 197: 221-231.
9. Tajbakhsh N, Suzuki K. Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs. *Pattern Recognition*, 2016, 63: 476-486.
10. Yang D, Zhang ST, Yan ZN, Tan CW, Li K, Metaxas D. Automated anatomical landmark detection on distal femur surface using convolutional neural network. In: *Proceedings of the 12th International Symposium on Biomedical Imaging*. New York, NY, USA: IEEE, 2015. 17-21.
11. de Vos BD, Wolterink JM, de Jong PA, Viergever MA, Išgum I. 2D image classification for 3D anatomy localization: employing deep convolutional neural networks. In: *Proceedings of the 9784, Medical Imaging 2016: Image Processing*. San Diego, California, US: SPIE, 2016, 9784: Article No. 97841Y.
12. Chen H, Yu LQ, Dou Q, Shi L, Mok VCT, Heng PA. Automatic detection of cerebral microbleeds via deep learning based 3D feature representation. In: *Proceedings of the 12th International Symposium on Biomedical Imaging (ISBI)*. New York, NY, USA: IEEE, 2015. 764-767.
13. Shin HC, Orton MR, Collins DJ, Doran SJ, Leach M O. Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35 (8): 1930-1943.
14. Cireşan DC, Giusti A, Gambardella LM, Schmidhuber J. Deep neural networks segment neuronal membranes in electron microscopy images. In: *Proceedings of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe, Nevada: Curran Associates Inc., 2012. 2843-2851.
15. Kumar N, Verma R, Sharma S, Bhargava S, Vahadane A, Sethi A. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Transactions on Medical Imaging*, 2017, 36 (7): 1550-1560.
16. Zhang WL, Li RJ, Deng HT, Wang L, Lin WL, Ji SW, Shen D G. Deep convolutional neural networks for multi-modality iso-intense infant brain image segmentation. *NeuroImage*, 2015, 108: 214-224.
17. Carneiro G, Nascimento JC, Freitas A. The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods. *IEEE Transactions on Image Processing*, 2012, 21 (3): 968-982.
18. Avendi MR, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Medical Image Analysis*, 2016, 30: 108-119.